GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 10, 2003, 15:59:18 ; Search time 6118 Seconds
(without alignments)
10912.800 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976
Perfect score: 1632
Sequence: 1 tcaggggaagaaaaatattc.....ttgtctggtgctcagattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 1316618

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30.6	1.9	31	6	AX248013	AX248013 Sequence
	2	30.6	1.9	31	6	AX248014	AX248014 Sequence
c	3	22	1.3	22	6	AX418163	AX418163 Sequence
	4	22	1.3	46	9	AY078814S1	AY078814 Hylobates
	5	22	1.3	46	9	AY078857S1	AY078857 Pan trogl
	6	22	1.3	46	9	AY078864S1	AY078864 Papio ham
	7	22	1.3	46	9	AY078892S1	AY078892 Pongo pyg
	8	22	1.3	46	9	AY078899S1	AY078899 Pongo pyg
	9	22	1.3	46	9	AY078906S1	AY078906 Gorilla g
	10	22	1.3	46	9	AY078914S1	AY078914 Pan trogl
	11	21.8	1.3	47	6	AX114333	AX114333 Sequence
	12	21	1.3	21	6	AX418162	AX418162 Sequence
	13	21	1.3	45	6	I60571	I60571 Sequence 25
	14	21	1.3	46	9	AY078828S1	AY078828 Pongo pyg
	15	21	1.3	46	9	AY078835S1	AY078835 Pongo pyg
	16	21	1.3	46	9	AY078842S1	AY078842 Pongo pyg
	17	20.6	1.3	21	6	AX094952	AX094952 Sequence
	18	20.6	1.3	21	6	AX094953	AX094953 Sequence
	19	20.6	1.3	21	6	AX094954	AX094954 Sequence
c	20	20.6	1.3	41	6	AR006790	AR006790 Sequence
c	21	20.6	1.3	41	6	AR135398	AR135398 Sequence
c	22	20.6	1.3	41	6	I71302	I71302 Sequence 40
c	23	20.6	1.3	47	6	AR289315	AR289315 Sequence
	24	20.6	1.3	50	6	AX685332	AX685332 Sequence
	25	20.4	1.2	31	6	A76877	A76877 Sequence 9
	26	20.4	1.2	35	11	C75834	C75834 Homo sapien
	27	20.4	1.2	46	9	AY078807S1	AY078807 Hylobates
	28	20.4	1.2	46	9	AY078821S1	AY078821 Hylobates
	29	20.4	1.2	46	9	AY078849S1	AY078849 Gorilla g
	30	20.4	1.2	46	9	AY078871S1	AY078871 Hylobates
	31	20.4	1.2	46	9	AY078878S1	AY078878 Hylobates
	32	20.4	1.2	46	9	AY078885S1	AY078885 Hylobates
c	33	20.2	1.2	50	6	AX685333	AX685333 Sequence
	34	20	1.2	41	6	AX517118	AX517118 Sequence
	35	20	1.2	41	6	AX519647	AX519647 Sequence
c	36	19.8	1.2	33	6	AR145735	AR145735 Sequence
c	37	19.8	1.2	41	6	AR006789	AR006789 Sequence
c	38	19.8	1.2	41	6	AR135397	AR135397 Sequence
c	39	19.8	1.2	41	6	I71301	I71301 Sequence 39
c	40	19.8	1.2	47	6	AR290337	AR290337 Sequence
	41	19.8	1.2	47	6	AX697662	AX697662 Sequence
	42	19.8	1.2	47	6	BD075576	BD075576 Secretory
	43	19.8	1.2	47	6	BD172436	BD172436 Secreted
	44	19.8	1.2	47	6	BD172755	BD172755 Secreted
	45	19.8	1.2	47	6	BD173074	BD173074 Secreted

ALIGNMENTS

RESULT 1
AX248013
LOCUS AX248013 31 bp DNA linear PAT 28-SEP-2001
DEFINITION Sequence 92 from Patent WO0166800.
ACCESSION AX248013
VERSION AX248013.1 GI:15862636
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cargill,M., Ireland,J.S. and Lander,E.S.
TITLE Human single nucleotide polymorphisms
JOURNAL Patent: WO 0166800-A 92 13-SEP-2001;

```
          WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT        5 a      10 c       8 g       7 t       1 others
ORIGIN

  Query Match             1.9%;  Score 30.6;  DB 6;  Length 31;
  Best Local Similarity   96.8%;  Pred. No. 7.3e+04;
  Matches   30;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        329 GGGACAAGCCTGCTCGCTGCCTTTGACCAGT 359
              ||||||||||||||:|||||||||||||||
Db          1 GGGACAAGCCTGCTCRCTGCCTTTGACCAGT 31


RESULT 2
AX248014
LOCUS       AX248014                  31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 93 from Patent WO0166800.
ACCESSION   AX248014
VERSION     AX248014.1  GI:15862637
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 93 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT        6 a      8 c       9 g       7 t       1 others
ORIGIN

  Query Match             1.9%;  Score 30.6;  DB 6;  Length 31;
  Best Local Similarity   96.8%;  Pred. No. 7.3e+04;
  Matches   30;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        973 GGTCACGGGCAGTGCCCATTGCACGTTTAAC 1003
              ||||||||||||||:|||||||||||||||
Db          1 GGTCACGGGCAGTGCSCATTGCACGTTTAAC 31


RESULT 3
AX418163/c
LOCUS       AX418163                  22 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 16 from Patent WO0202620.
ACCESSION   AX418163
VERSION     AX418163.1  GI:21523173
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Lennon,V.A., Yu,Z., Kryzer,T.J. and Griesmann,G.E.
  TITLE     Crmp-5 (collapsin response-mediator protein) encoding nucleic acid,
            polypeptide and uses thereof
  JOURNAL   Patent: WO 0202620-A 16 10-JAN-2002;
            MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide for PCR"
BASE COUNT        2 a      4 c       8 g       8 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+06;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1229 GACCCCGACAGCGTTAAAACCA 1250
              ||||||||||||||||||||||
Db         22 GACCCCGACAGCGTTAAAACCA 1


RESULT 4
AY078814S1
LOCUS       AY078814S1                46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Hylobates syndactylus isolate B1533 putative CD209L1 protein
            (CD209L1) gene, exon 1.
ACCESSION   AY078814
VERSION     AY078814.1  GI:27356793
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Hylobates syndactylus (siamang)
  ORGANISM  Hylobates syndactylus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A. and Carrington,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Hylobates syndactylus"
                     /mol_type="genomic DNA"
                     /isolate="B1533"
                     /db_xref="taxon:9590"
     exon            <1. .46
                     /gene="CD209L1"
                     /note="similar to Homo sapiens CD209L1 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       10 a      13 c      16 g       7 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||  |  ||||||||||||||||  ||    ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGGGTGCAGCCGCTGGG 38


RESULT 5
AY078857S1
LOCUS       AY078857S1                46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Pan troglodytes isolate B1437 putative CD209L1 protein (CD209L1)
            gene, exon 1.
ACCESSION   AY078857
VERSION     AY078857.1  GI:27356847
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
```

```
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A. and Carrington,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="B1437"
                     /db_xref="taxon:9598"
     exon            <1. .46
                     /gene="CD209L1"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       11 a     12 c     16 g      7 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||.|  |||||||||||||||  ||    ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGGGTGCAGCAGCTGGG 38


RESULT 6
AY078864S1
LOCUS       AY078864S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Papio hamadryas isolate B854 putative CD209 protein (CD209) gene,
            exon 1.
ACCESSION   AY078864
VERSION     AY078864.1  GI:27356858
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Papio hamadryas"
                     /mol_type="genomic DNA"
                     /isolate="B854"
                     /db_xref="taxon:9557"
     exon            <1. .46
                     /gene="CD209"
                     /note="similar to Homo sapiens CD209 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       12 a     13 c     14 g      7 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||.|  |||||||||||||||  ||    ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGACTGCAGCAGCTGGG 38


RESULT 7
AY078892S1
LOCUS       AY078892S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Pongo pygmaeus isolate Ppy21 putative CD209 protein (CD209) gene,
            exon 1.
ACCESSION   AY078892
VERSION     AY078892.1  GI:27356894
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="Ppy21"
                     /db_xref="taxon:9600"
     exon            <1. .46
                     /gene="CD209"
                     /note="similar to Homo sapiens CD209 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       12 a     12 c     14 g      8 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||.|  |||||||||||||||  ||    ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGACTGCAGCAGCTGGG 38


RESULT 8
AY078899S1
LOCUS       AY078899S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Pongo pygmaeus isolate Ppy91 putative CD209 protein (CD209) gene,
            exon 1.
ACCESSION   AY078899
VERSION     AY078899.1  GI:27356903
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
```

```
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="Ppy91"
                     /db_xref="taxon:9600"
     exon            <1. .46
                     /gene="CD209"
                     /note="similar to Homo sapiens CD209 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       12 a     12 c     14 g      8 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||.|  |||||||||||||| | ||   ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGACTGCAGCAGCTGGG 38


RESULT 9
AY078906S1
LOCUS       AY078906S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Gorilla gorilla isolate Ggo3 putative CD209 protein (CD209) gene,
            exon 1.
ACCESSION   AY078906
VERSION     AY078906.1  GI:27356912
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /isolate="Ggo3"
                     /db_xref="taxon:9593"
     exon            <1. .46
                     /gene="CD209"
                     /note="similar to Homo sapiens CD209 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       12 a     13 c     14 g      7 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||.|  |||||||||||||| | ||   ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGACTGCAGCAGCTGGG 38


RESULT 10
AY078914S1
LOCUS       AY078914S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Pan troglodytes isolate B1437 putative CD209 protein (CD209) gene,
            exon 1.
ACCESSION   AY078914
VERSION     AY078914.1  GI:27356921
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /isolate="B1437"
                     /db_xref="taxon:9598"
     exon            <1. .46
                     /gene="CD209"
                     /note="similar to Homo sapiens CD209 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       12 a     13 c     14 g      7 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 9;  Length 46;
  Best Local Similarity   73.7%;  Pred. No. 5.1e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGGG 289
              ||||.|  |||||||||||||| | ||   ||||||
Db          1 ATGAGTGACTCCAAGGAACCAAGACTGCAGCAGCTGGG 38


RESULT 11
AX114333
LOCUS       AX114333                 47 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 2 from Patent WO0129257.
ACCESSION   AX114333
VERSION     AX114333.1  GI:14031297
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schork,N. and Skierczynski,B.
  TITLE     Methods of genetic cluster analysis and use thereof
  JOURNAL   Patent: WO 0129257-A 2 26-APR-2001;
            GENSET (FR)
FEATURES             Location/Qualifiers
```

```
     source          1. .47
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     variation       24
                     /note="10-24-234 : polymorphic base A or G"
BASE COUNT       11 a     10 c     18 g      7 t      1 others
ORIGIN

  Query Match             1.3%;  Score 21.8;  DB 6;  Length 47;
  Best Local Similarity   67.4%;  Pred. No. 5.6e+06;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        425 CATAAGGGCATCCAGGAGGAGATGGAAGCGCTTGTGAAGGATC 467
              ||  || ||   | | ||||| |: ||| |||| ||||||| |
Db          1 CAGGAGCGCTGTCGGCAGGAGGTRCAAGAGCTTCTGAAGGACC 43


RESULT 12
AX418162
LOCUS       AX418162                 21 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 15 from Patent WO0202620.
ACCESSION   AX418162
VERSION     AX418162.1  GI:21523172
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Lennon,V.A., Yu,Z., Kryzer,T.J. and Griesmann,G.E.
  TITLE     Crmp-5 (collapsin response-mediator protein) encoding nucleic acid,
            polypeptide and uses thereof
  JOURNAL   Patent: WO 0202620-A 15 10-JAN-2002;
            MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide for PCR"
BASE COUNT        6 a      6 c      5 g      4 t
ORIGIN

  Query Match             1.3%;  Score 21;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 8.9e+06;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        505 AGATCGCTTCCAGCTAACGGA 525
              |||||||||||||||||||||
Db          1 AGATCGCTTCCAGCTAACGGA 21


RESULT 13
I60571
LOCUS       I60571                   45 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 25 from patent US 5656725.
ACCESSION   I60571
VERSION     I60571.1  GI:2479016
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Chittenden,T.D. and Lutz,R.J.
  TITLE     Peptides and compositions which modulate apoptosis
  JOURNAL   Patent: US 5656725-A 25 12-AUG-1997;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT       12 a     11 c     15 g      7 t
ORIGIN

  Query Match             1.3%;  Score 21;  DB 6;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 8.3e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | |||||||| |  ||| |||||  ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38


RESULT 14
AY078828S1
LOCUS       AY078828S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Pongo pygmaeus isolate Ppy21 putative CD209L1 protein (CD209L1)
            gene, exon 1.
ACCESSION   AY078828
VERSION     AY078828.1  GI:27356811
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A. and Carrington,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="Ppy21"
                     /db_xref="taxon:9600"
     exon            <1. .46
                     /gene="CD209L1"
                     /note="similar to Homo sapiens CD209L1 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       13 a     12 c     14 g      7 t
ORIGIN

  Query Match             1.3%;  Score 21;  DB 9;  Length 46;
  Best Local Similarity   73.0%;  Pred. No. 8.3e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGG 288
              |||| |   |||||||||||||||  ||    |||||
Db          1 ATGAGTGACTCCAAGGAACCAAGGGTGCAGCAGCTGG 37


RESULT 15
AY078835S1
LOCUS       AY078835S1               46 bp    DNA     linear   PRI 21-DEC-2002
DEFINITION  Pongo pygmaeus isolate Ppy26 putative CD209L1 protein (CD209L1)
            gene, exon 1.
ACCESSION   AY078835
VERSION     AY078835.1  GI:27356820
KEYWORDS    .
SEGMENT     1 of 7
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P.,
            Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and
            Carrington,M.
```

```
  TITLE     Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
  JOURNAL   J. Virol. 77 (1), 217-227 (2003)
   PUBMED   12477827
REFERENCE   2  (bases 1 to 46)
  AUTHORS   Bashirova,A.A. and Carrington,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
            NCI-Frederick, Frederick, MD 21702, USA
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="Ppy26"
                     /db_xref="taxon:9600"
     exon            <1. .46
                     /gene="CD209L1"
                     /note="similar to Homo sapiens CD209L1 gene"
                     /number=1
                     /evidence=not_experimental
BASE COUNT       13 a     12 c     14 g      7 t
ORIGIN

  Query Match             1.3%;  Score 21;  DB 9;  Length 46;
  Best Local Similarity   73.0%;  Pred. No. 8.3e+06;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        252 ATGATTTCTTCCAAGGAACCAAGGCGGCCCTGGCTGG 288
              |||| |   ||||||||||||||||  ||    |||||
Db          1 ATGAGTGACTCCAAGGAACCAAGGGTGCAGCAGCTGG 37


Search completed: December 10, 2003, 21:37:37
Job time : 6122 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 10, 2003, 18:45:12 ; Search time 3660 Seconds
(without alignments)
10837.406 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976
Perfect score: 1632
Sequence: 1 tcaggggaagaaaaatattc.....ttgtctggtgctcagattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 124404

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		26.6	1.6	43	9	AA975126	AA975126 on06e08.s
c 2		22.4	1.4	50	9	AU106847	AU106847 AU106847
c 3		20.6	1.3	50	9	AU106844	AU106844 AU106844
c 4		20.6	1.3	50	9	AU106846	AU106846 AU106846
c 5		20.6	1.3	50	9	AU106848	AU106848 AU106848
c 6		20.4	1.2	41	28	AZ647118	AZ647118 1M0513O23
7		20.4	1.2	44	28	AZ983982	AZ983982 2M0265O13
c 8		20.2	1.2	44	28	AZ413456	AZ413456 1M0197E14
9		20.2	1.2	45	14	H67715	H67715 yr72c07.s1
10		19.8	1.2	50	9	AU103525	AU103525 AU103525
c 11		19.6	1.2	42	28	AZ633442	AZ633442 1M0488N08
12		19.6	1.2	43	28	BH234824	BH234824 MEAA_E02.
c 13		19.6	1.2	49	28	AZ964788	AZ964788 2M0234C11
14		19.6	1.2	50	9	AU106256	AU106256 AU106256
15		19.4	1.2	40	14	H25033	H25033 yl39f09.r1
16		19.4	1.2	42	9	AL796248	AL796248 AL796248
17		19.4	1.2	50	9	AU102320	AU102320 AU102320
18		19.2	1.2	49	29	BX133642	BX133642 Danio rer
c 19		19.2	1.2	50	9	AU103396	AU103396 AU103396
20		19.2	1.2	50	9	AU105960	AU105960 AU105960
21		19.2	1.2	50	9	AU107563	AU107563 AU107563
c 22		19	1.2	47	28	BH640414	BH640414 1008035F1
23		19	1.2	48	28	BH910594	BH910594 SALK_0605
24		19	1.2	50	9	AU107084	AU107084 AU107084
c 25		18.8	1.2	45	29	AL940299	AL940299 Arabidops
c 26		18.8	1.2	48	12	BI753336	BI753336 603026191
c 27		18.8	1.2	49	9	AA151661	AA151661 zo29a02.s
28		18.8	1.2	49	14	H22280	H22280 yl36b03.r1
29		18.8	1.2	50	9	AU103300	AU103300 AU103300
30		18.8	1.2	50	9	AU107374	AU107374 AU107374
c 31		18.8	1.2	50	9	AU107543	AU107543 AU107543
c 32		18.8	1.2	50	28	AZ920484	AZ920484 1006020A0
33		18.6	1.1	37	9	AI721602	AI721602 fc29h09.x
c 34		18.6	1.1	43	9	AA676294	AA676294 ad37d06.s
c 35		18.6	1.1	49	9	AL776769	AL776769 AL776769
36		18.4	1.1	50	9	AU106434	AU106434 AU106434
37		18.4	1.1	50	9	AU106786	AU106786 AU106786
38		18.4	1.1	50	9	AU106790	AU106790 AU106790
39		18.4	1.1	50	9	AU106794	AU106794 AU106794
40		18.4	1.1	50	9	AU106795	AU106795 AU106795
c 41		18.4	1.1	50	9	AU107060	AU107060 AU107060
c 42		18.2	1.1	34	28	AZ663610	AZ663610 1M0543O16
c 43		18.2	1.1	43	9	AV833356	AV833356 AV833356
c 44		18.2	1.1	46	28	BH865380	BH865380 SALK_0983
c 45		18.2	1.1	47	10	BF969163	BF969163 602269853

ALIGNMENTS

RESULT 1
AA975126
LOCUS AA975126 43 bp mRNA linear EST 23-OCT-1998
DEFINITION on06e08.s1 NCI_CGAP_Lei2 Homo sapiens cDNA clone IMAGE:1555910 3' similar to TR:Q13024 Q13024 COLLAPSIN RESPONSE MEDIATOR PROTEIN CRMP-1. [1] ;, mRNA sequence.
ACCESSION AA975126
VERSION AA975126.1 GI:3150918
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 43)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
unknown library type
Trace considered overall poor quality
Insert Length: 715 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .43

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1555910"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lei2"
                     /note="Organ: soft tissue; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-AACTGGAAGAATTCGCGGCCGCAATCGTTTTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT        8 a      11 c      15 g       9 t
ORIGIN

  Query Match             1.6%;  Score 26.6;  DB 9;  Length 43;
  Best Local Similarity   87.9%;  Pred. No. 1.3e+04;
  Matches   29;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1278 TCGAGTACAACATCTTTGAAGGCATGGAGTGCC 1310
              | ||||||||||||||| || || ||||||||||
Db          3 TGGAGTACAACATCTTCGAGGGTATGGAGTGCC 35


RESULT 2
AU106847/c
LOCUS       AU106847                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106847 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI04566, mRNA sequence.
ACCESSION   AU106847
VERSION     AU106847.1  GI:13556368
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI04566"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       11 a      10 c      15 g      14 t
ORIGIN

  Query Match             1.4%;  Score 22.4;  DB 9;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 2e+05;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        816 ATGGCGAGCCCATCACTGCCAGCTTGGGAACGGACGGCTCCCATTACT 863
              ||||||| | ||  ||  |   ||||| || |  |||||| ||| | |

Db         49 ATGGCGATCACACAACCACTCCCTTGGAAAGGCTCGGCTCACATAATT 2


RESULT 3
AU106844/c
LOCUS       AU106844                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106844 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC02437, mRNA sequence.
ACCESSION   AU106844
VERSION     AU106844.1  GI:13556365
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC02437"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        8 a      13 c      14 g      15 t
ORIGIN

  Query Match             1.3%;  Score 20.6;  DB 9;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 6.2e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        816 ATGGCGAGCCCATCACTGCCAGCTTGGGAACGGACGGCTCCCA 858
              ||||||| | ||  ||  |   ||||| || |  |||||| ||
Db         46 ATGGCGATCACACAACCACTCCCTTGGAAAGGTGCGGCTCACA 4


RESULT 4
AU106846/c
LOCUS       AU106846                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106846 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI00647, mRNA sequence.
ACCESSION   AU106846
VERSION     AU106846.1  GI:13556367
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
```

```
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
          ,S. Construction and characterization of a full length-enriched and
          a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI00647"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        9 a     11 c     16 g     14 t
ORIGIN

  Query Match             1.3%;  Score 20.6;  DB 9;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 6.2e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        816 ATGGCGAGCCCATCACTGCCAGCTTGGGAACGGACGGCTCCCA 858
              ||||||| | ||  ||  |  ||||| || |  |||||| ||
Db         44 ATGGCGATCACACAACCACTCCCTTGGAAAGGTTCGGCTCACA 2


RESULT 5
AU106848/c
LOCUS       AU106848                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106848 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI05259, mRNA sequence.
ACCESSION   AU106848
VERSION     AU106848.1  GI:13556369
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05259"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        8 a     13 c     14 g     15 t
ORIGIN

  Query Match             1.3%;  Score 20.6;  DB 9;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 6.2e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        816 ATGGCGAGCCCATCACTGCCAGCTTGGGAACGGACGGCTCCCA 858
              ||||||| | ||  ||  |  ||||| || |  |||||| ||
Db         46 ATGGCGATCACACAACCACTCCCTTGGAAAGGTGCGGCTCACA 4


RESULT 6
AZ647118/c
LOCUS       AZ647118                 41 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0513O23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0513O23 F, genomic survey sequence.
ACCESSION   AZ647118
VERSION     AZ647118.1  GI:11778266
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0513  row: O  column: 23
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop:. 41.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0513O23"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a     15 c     13 g      4 t
ORIGIN

  Query Match             1.2%;  Score 20.4;  DB 28;  Length 41;
  Best Local Similarity   80.0%;  Pred. No. 6.4e+05;
  Matches   24;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        338 CTGCTCGCTGCCTTTGACCAGTGGAGGGAA 367
              ||||| ||||||| ||  ||||| ||| ||
Db         31 CTGCTGGCTGCCTGTGGTCAGTGCAGGCAA 2
```

```
RESULT 7
AZ983982
LOCUS       AZ983982                 44 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0265O13F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0265O13 F, genomic survey sequence.
ACCESSION   AZ983982
VERSION     AZ983982.1  GI:13855209
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0265  row: O  column: 13
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 44.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0265O13"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       17 a      2 c     21 g      4 t
ORIGIN

  Query Match             1.2%;  Score 20.4;  DB 28;  Length 44;
  Best Local Similarity   71.1%;  Pred. No. 6.6e+05;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        428 AAGGGCATCCAGGAGGAGATGGAAGCGCTTGTGAAGGA 465
              ||  | |||  ||||| ||||||||  |   || |||||
Db          4 AAATGAATCGCGGAGGGGATGGAAGAGAGGGTAAAGGA 41


RESULT 8
AZ413456/c
LOCUS       AZ413456                 44 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0197E14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0197E14 R, genomic survey sequence.
ACCESSION   AZ413456
VERSION     AZ413456.1  GI:10537469
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0197  row: E  column: 14
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 44.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0197E14"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       8 a      3 c     26 g      7 t
ORIGIN

  Query Match             1.2%;  Score 20.2;  DB 28;  Length 44;
  Best Local Similarity   68.3%;  Pred. No. 7.5e+05;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        894 TTGTCACCTCCCCACCCTTGAGCCCTGATCCAACCACTCCA 934
```

```
                 | ||| || |||| || ||   ||  | ||| ||||||||
Db         43 TAGTCCCCCCCCCCACCCTGCTTCCCTACCCATCCACTCCA 3


RESULT 9
H67715
LOCUS       H67715                   45 bp    mRNA    linear   EST 18-OCT-1995
DEFINITION  yr72c07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:210828 3' similar to gb:J00314_rna2 TUBULIN BETA-1 CHAIN
            (HUMAN);, mRNA sequence.
ACCESSION   H67715
VERSION     H67715.1  GI:1026455
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 817
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 817   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3783669"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:210828"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        6 a      20 c       8 g       9 t       2 others
ORIGIN

  Query Match             1.2%;  Score 20.2;  DB 14;  Length 45;
  Best Local Similarity   71.4%;  Pred. No. 7.6e+05;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        876 GGGCCAAGGCTGCTGCCTTTGTCACCTCCCCACCC 910
              || ||||| ||| |  ||||  | ||||||  |||
Db          9 GGACCAAGNCTGGTCTCTTTCCCCCCTCCCANCCC 43


RESULT 10
AU103525
LOCUS       AU103525                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU103525 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC05285, mRNA sequence.
ACCESSION   AU103525
VERSION     AU103525.1  GI:13553046
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC05285"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       13 a      15 c      16 g       6 t
ORIGIN

  Query Match             1.2%;  Score 19.8;  DB 9;  Length 50;
  Best Local Similarity   63.8%;  Pred. No. 1e+06;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1122 AGAACCAGTTTGTGGCTGTGACCAGCACCAATGCAGCCAAAGTCTTC 1168
              || | ||| ||| ||| | |  |||||||   |   ||| | || ||
Db          1 AGCAGCAGGTTGAGGCAGAGCACAGCACCGTCGGGACCAGACTCGTC 47


RESULT 11
AZ633442/c
LOCUS       AZ633442                 42 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0488N08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0488N08 R, genomic survey sequence.
ACCESSION   AZ633442
VERSION     AZ633442.1  GI:11755632
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
```

```
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0488  row: N  column: 08
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 42.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0488N08"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT        0 a     20 c      6 g     16 t
ORIGIN

  Query Match             1.2%;  Score 19.6;  DB 28;  Length 42;
  Best Local Similarity   73.5%;  Pred. No. 1.1e+06;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        416 AGCGAGTGGCATAAGGGCATCCAGGAGGAGATGG 449
              ||| || || | ||||| | | ||||| ||| ||
Db         38 AGCCAGGGGAAGAAGGGAAGCAAGGAGCAGAGGG 5


RESULT 12
BH234824
LOCUS       BH234824                  43 bp    DNA     linear   GSS 01-JAN-2002
DEFINITION  MEAA_E02.x E Spiroplasma kunkelii genomic clone MEAA_E02.x, genomic
            survey sequence.
ACCESSION   BH234824
VERSION     BH234824.1  GI:18030292
KEYWORDS    GSS.
SOURCE      Spiroplasma kunkelii
  ORGANISM  Spiroplasma kunkelii
            Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
            Spiroplasmataceae; Spiroplasma.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Hogenhout,S.A.
  TITLE     Genomic sequences from Spiroplasma kunkelii strain M2
  JOURNAL   Unpublished
COMMENT     Contact: Hogenhout SA
            Department of Entomology
            The Ohio State University-OARDC
            120 Thorne Hall, 1680 Madison Avenue, Wooster, OH 44691, USA
            Tel: 330 263 3730
            Fax: 330 263 3686
            Email: hogenhout.1@osu.edu
            Plate: AA  row: E  column: 02
            Class: EcoRI fragments.
FEATURES              Location/Qualifiers
     source           1. .43
                      /organism="Spiroplasma kunkelii"
                      /mol_type="genomic DNA"
                      /strain="M2"
                      /db_xref="taxon:47834"
                      /clone="MEAA_E02.x"
                      /clone_lib="E"
BASE COUNT       14 a      7 c      3 g     19 t
ORIGIN

  Query Match             1.2%;  Score 19.6;  DB 28;  Length 43;
  Best Local Similarity   73.5%;  Pred. No. 1.1e+06;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        475 AAATTCCTTCCTCGTGTACATGGCTTTCAAAGAT 508
              ||||| |||| || | |   |  ||||||||||
Db          3 AAATTTCTTCTTCTTATTTTTCTCTTTCAAAGAT 36


RESULT 13
AZ964788/c
LOCUS       AZ964788                  49 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0234C11R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0234C11 R, genomic survey sequence.
ACCESSION   AZ964788
VERSION     AZ964788.1  GI:13836015
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0234  row: C  column: 11
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 49.
FEATURES              Location/Qualifiers
     source           1. .49
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0234C11"
                      /sex="Female"
                      /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC2M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (female) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
```

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```
BASE COUNT       12 a      6 c     24 g      7 t
ORIGIN

  Query Match             1.2%;  Score 19.6;  DB 28;  Length 49;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+06;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        930 CTCCAGACTTTCTCAACTCCTTGCTGTCCTGTGGAGACCTCC 971
              |||||| | ||   | |||  | |||||||    || |||||
Db         49 CTCCAGCCATTAGGACCTCTCTCCTGTCCTCCACAGTCCTCC 8


RESULT 14
AU106256
LOCUS       AU106256                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106256 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            KAIA4033, mRNA sequence.
ACCESSION   AU106256
VERSION     AU106256.1  GI:13555777
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAIA4033"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       14 a     16 c      9 g     11 t
ORIGIN

  Query Match             1.2%;  Score 19.6;  DB 9;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+06;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        968 CTCCAGGTCACGGGCAGTGCCCATTGCACGTTTAACACTGCC 1009
              | ||||| || || ||  ||    || || ||||| ||| ||
Db          8 CCCCAGGGCAAGGTCACAGCTGCCTGAACTTTTAAAACTCCC 49
```

```
RESULT 15
H25033
LOCUS       H25033                    40 bp    mRNA    linear   EST 07-JUL-1995
DEFINITION  yl39f09.r1 Soares breast 3NbHBst Homo sapiens cDNA clone
            IMAGE:160649 5' similar to SP:S32603 S32603 COLLAGEN ALPHA 1(VI)
            CHAIN - MOUSE ;, mRNA sequence.
ACCESSION   H25033
VERSION     H25033.1  GI:893932
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1426
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1426   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:574710"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:160649"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares breast 3NbHBst"
                     /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of a modified pT7T3 vector (Pharmacia).
                     Library went through one round of normalization to a Cot =
                     20. Library constructed by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT        8 a      9 c     14 g      7 t      2 others
ORIGIN

  Query Match             1.2%;  Score 19.4;  DB 14;  Length 40;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+06;
  Matches   26;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        664 GCTGAGCCGACCTGAGGAGGTCGAGGCCGAAGCCGTGAA 702
              |||||| ||| |||||   ||| ||  |||  | | | |
Db          1 GCTGAGNCGATCTGAGCTNGTCAAGTTCGAGCCAGGGCA 39


Search completed: December 10, 2003, 22:38:55
Job time : 3666 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 10, 2003, 21:37:47 ; Search time 548 Seconds
(without alignments)
9898.018 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976
Perfect score: 1632
Sequence: 1 tcaggggaagaaaaatattc.....ttgtctggtgctcagattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 1223696

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30.6	1.9	31	9	US-09-801-274-92	Sequence 92, Appl
	2	30.6	1.9	31	9	US-09-801-274-93	Sequence 93, Appl
c	3	20.6	1.3	41	10	US-09-850-165-64	Sequence 64, Appl
c	4	20.6	1.3	41	15	US-10-211-357-33	Sequence 33, Appl
c	5	20.6	1.3	45	13	US-10-282-960-95	Sequence 95, Appl
	6	20.4	1.2	37	13	US-10-150-048-13	Sequence 13, Appl
c	7	20.4	1.2	40	11	US-09-865-281A-6	Sequence 6, Appli
	8	20.4	1.2	41	13	US-10-224-683-67	Sequence 67, Appl
	9	20.2	1.2	37	13	US-10-027-632-52319	Sequence 52319, A
	10	20.2	1.2	37	13	US-10-027-632-52331	Sequence 52331, A
	11	20.2	1.2	37	14	US-10-027-632-52319	Sequence 52319, A
	12	20.2	1.2	37	14	US-10-027-632-52331	Sequence 52331, A
c	13	20	1.2	20	15	US-10-006-911-34	Sequence 34, Appl
c	14	20	1.2	20	15	US-10-006-911-35	Sequence 35, Appl
c	15	20	1.2	20	15	US-10-006-911-36	Sequence 36, Appl
c	16	20	1.2	20	15	US-10-006-911-37	Sequence 37, Appl
c	17	20	1.2	20	15	US-10-006-911-38	Sequence 38, Appl
c	18	20	1.2	20	15	US-10-006-911-39	Sequence 39, Appl
c	19	20	1.2	20	15	US-10-006-911-40	Sequence 40, Appl
c	20	20	1.2	20	15	US-10-006-911-41	Sequence 41, Appl
c	21	20	1.2	20	15	US-10-006-911-42	Sequence 42, Appl
c	22	20	1.2	20	15	US-10-006-911-43	Sequence 43, Appl
c	23	20	1.2	20	15	US-10-006-911-44	Sequence 44, Appl
c	24	20	1.2	20	15	US-10-006-911-45	Sequence 45, Appl
c	25	20	1.2	20	15	US-10-006-911-46	Sequence 46, Appl
c	26	20	1.2	20	15	US-10-006-911-47	Sequence 47, Appl
c	27	20	1.2	20	15	US-10-006-911-48	Sequence 48, Appl
c	28	20	1.2	20	15	US-10-006-911-49	Sequence 49, Appl
c	29	20	1.2	20	15	US-10-006-911-50	Sequence 50, Appl
c	30	20	1.2	20	15	US-10-006-911-51	Sequence 51, Appl
c	31	20	1.2	20	15	US-10-006-911-52	Sequence 52, Appl
c	32	20	1.2	20	15	US-10-006-911-53	Sequence 53, Appl
c	33	20	1.2	20	15	US-10-006-911-54	Sequence 54, Appl
c	34	20	1.2	20	15	US-10-006-911-55	Sequence 55, Appl
c	35	20	1.2	20	15	US-10-006-911-56	Sequence 56, Appl
c	36	20	1.2	20	15	US-10-006-911-57	Sequence 57, Appl
c	37	20	1.2	20	15	US-10-006-911-58	Sequence 58, Appl
c	38	20	1.2	20	15	US-10-006-911-59	Sequence 59, Appl
c	39	20	1.2	20	15	US-10-006-911-60	Sequence 60, Appl
c	40	20	1.2	20	15	US-10-006-911-61	Sequence 61, Appl
c	41	20	1.2	20	15	US-10-006-911-62	Sequence 62, Appl
c	42	20	1.2	20	15	US-10-006-911-63	Sequence 63, Appl
c	43	20	1.2	20	15	US-10-006-911-64	Sequence 64, Appl
c	44	20	1.2	50	16	US-10-179-730-7	Sequence 7, Appli
c	45	19.8	1.2	41	10	US-09-850-165-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-801-274-92
; Sequence 92, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-92

Query Match 1.9%; Score 30.6; DB 9; Length 31;
Best Local Similarity 96.8%; Pred. No. 16;
Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy        329 GGGACAAGCCTGCTCGCTGCCTTTGACCAGT 359
              |||||||||||||||:|||||||||||||||
Db          1 GGGACAAGCCTGCTCRCTGCCTTTGACCAGT 31
```

RESULT 2
US-09-801-274-93
; Sequence 93, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:

```
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 93
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-93

  Query Match             1.9%;  Score 30.6;  DB 9;  Length 31;
  Best Local Similarity   96.8%;  Pred. No. 16;
  Matches   30;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        973 GGTCACGGGCAGTGCCCATTGCACGTTTAAC 1003
              |||||||||||||||:|||||||||||||||
Db          1 GGTCACGGGCAGTGCSCATTGCACGTTTAAC 31


RESULT 3
US-09-850-165-64/c
; Sequence 64, Application US/09850165
; Patent No. US20020150580A1
; GENERAL INFORMATION:
;  APPLICANT: NEWMAN, ROLAND A.
;  APPLICANT:  HANNA, NABIL
;  APPLICANT:  RAAB, RONALD W.
;  TITLE OF INVENTION: RECOMBINANT ANTIBODIES FOR HUMAN THERAPY
;  FILE REFERENCE: 037003-0280614
;  CURRENT APPLICATION NUMBER: US/09/850,165
;  CURRENT FILING DATE:  2001-05-08
;  PRIOR APPLICATION NUMBER: 09/082,472
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: 08/476,237
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: 08/397,072
;  PRIOR FILING DATE: 1995-04-17
;  PRIOR APPLICATION NUMBER: 07/912,292
;  PRIOR FILING DATE: 1992-07-10
;  PRIOR APPLICATION NUMBER: 07/856,281
;  PRIOR FILING DATE: 1992-03-23
;  PRIOR APPLICATION NUMBER: 07/735,064
;  PRIOR FILING DATE: 1991-07-25
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-850-165-64

  Query Match             1.3%;  Score 20.6;  DB 10;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 2.9e+04;
  Matches   27;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
              ||||| ||||  |||:||| ||    |||  || |||||
Db         41 CTGAGTGGGGGCCCTBGTGTCCATGGTGAGAGATCTGTG 3


RESULT 4
US-10-211-357-33/c
; Sequence 33, Application US/10211357
; Publication No. US20030077275A1
;  GENERAL INFORMATION:
;       APPLICANT: Hanna, Nabil
;                  Newman, Roland A.
;                  Reff, Mitchell E.
;       TITLE OF INVENTION: Recombinant Anti-CD4 Antibodies for Human
;                           Therapy
;       NUMBER OF SEQUENCES: 59
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
;            STREET: 699 Prince Street
;            CITY: Alexandria
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22314-3187
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/211,357
;            FILING DATE: 05-Aug-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/612,914A
;            FILING DATE: 10-Jul-2000
;            APPLICATION NUMBER: US 08/523,894
;            FILING DATE: 06-SEP-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Teskin, Robin L.
;            REGISTRATION NUMBER: 35,030
;            REFERENCE/DOCKET NUMBER: 012712-165
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-836-6620
;            TELEFAX: 703-836-2021
;  INFORMATION FOR SEQ ID NO: 33:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 41 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: Human or Monkey
;       POSITION IN GENOME:
;            CHROMOSOME/SEGMENT: kappa light chain primer with Bgl II site
;       SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-10-211-357-33

  Query Match             1.3%;  Score 20.6;  DB 15;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 2.9e+04;
  Matches   27;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
              ||||| ||||  |||:||| ||    |||  || |||||
Db         41 CTGAGTGGGGGCCCTBGTGTCCATGGTGAGAGATCTGTG 3


RESULT 5
US-10-282-960-95/c
; Sequence 95, Application US/10282960
; Publication No. US20030143228A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Si-Yi
;  APPLICANT:  Zhaoyang, You
;  APPLICANT:  Schroers, Roland
;  TITLE OF INVENTION: Human Telomerase Reverse Transcriptase as a Class-II Restricted
;  TITLE OF INVENTION:  Associated Antigen
;  FILE REFERENCE: P02193US1
;  CURRENT APPLICATION NUMBER: US/10/282,960
```

```
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US 60/345,012
;  PRIOR FILING DATE: 2001-10-29
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Human
US-10-282-960-95

  Query Match             1.3%;  Score 20.6;  DB 13;  Length 45;
  Best Local Similarity   67.4%;  Pred. No. 3.1e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        810 TGGTGTATGGCGAGCCCATCACTGCCAGCTTGGGAACGGACGG 852
              ||||||    || |||| ||||  ||| ||||  || | || |
Db         43 TGGTGTCGTACGCGCCCGTCACATCCACCTTGACAAAGTACAG 1


RESULT 6
US-10-150-048-13
; Sequence 13, Application US/10150048
; Publication No. US20030215907A1
; GENERAL INFORMATION:
;  APPLICANT: Samuelson, James
;  APPLICANT:  Xu, Shuang-yong
;  TITLE OF INVENTION: Method For Cloning And Expression of PpuMI Restriction Endonuclea
;  TITLE OF INVENTION:  PpuMI Methylase in E. coli
;  FILE REFERENCE: NEB-204
;  CURRENT APPLICATION NUMBER: US/10/150,048
;  CURRENT FILING DATE:  2002-05-17
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Primer
US-10-150-048-13

  Query Match             1.2%;  Score 20.4;  DB 13;  Length 37;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+04;
  Matches   24;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        412 CATCAGCGAGTGGCATAAGGGCATCCAGGA 441
              || || |   ||||| ||||||||||||||
Db          4 CACCACCATATGGCAAAAGGGCATCCAGGA 33


RESULT 7
US-09-865-281A-6/c
; Sequence 6, Application US/09865281A
; Publication No. US20030103984A1
; GENERAL INFORMATION:
;  APPLICANT: Kohler, Heinz
;  TITLE OF INVENTION: FUSION PROTEINS OF BIOLOGICALLY ACTIVE PEPTIDES AND ANTIBODIES
;  FILE REFERENCE: 411.35629PC2
;  CURRENT APPLICATION NUMBER: US/09/865,281A
;  CURRENT FILING DATE:  2001-05-29
;  PRIOR APPLICATION NUMBER: 09/070,907
;  PRIOR FILING DATE: 1998-05-04
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(40)
;   OTHER INFORMATION: Synthesized oligonucleotide primer for Stat1 SH2 cDNA
US-09-865-281A-6

  Query Match             1.2%;  Score 20.4;  DB 11;  Length 40;
  Best Local Similarity   71.1%;  Pred. No. 3.4e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        601 CATCATTGCAGAGGAGCAGCAGAGGATCCTGGATCTGG 638
              ||||||| |||||  | |||||  | |   |||||  ||
Db         38 CATCATTCCAGAGAGGGAGCAGGTGTTTGGGGATCCGG 1


RESULT 8
US-10-224-683-67
; Sequence 67, Application US/10224683
; Publication No. US20030162192A1
; GENERAL INFORMATION:
;  APPLICANT: Sotos, John
;  APPLICANT:  Rienhoff, Jr., Hugh
;  APPLICANT:  Guida, Marco
;  APPLICANT:  Curran, Mark
;  TITLE OF INVENTION: Polymorphisms Associated with Ion-Channel Disease
;  FILE REFERENCE: 4389-33
;  CURRENT APPLICATION NUMBER: US/10/224,683
;  CURRENT FILING DATE:  2002-01-06
;  PRIOR APPLICATION NUMBER: 60/314,331
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/378,521
;  PRIOR FILING DATE: 2002-05-06
;  NUMBER OF SEQ ID NOS: 185
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-224-683-67

  Query Match             1.2%;  Score 20.4;  DB 13;  Length 41;
  Best Local Similarity   71.1%;  Pred. No. 3.4e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        419 GAGTGGCATAAGGGCATCCAGGAGGAGATGGAAGCGCT 456
              |||  | | ||| || | |||||||  ||||||  |||
Db          4 GAGAAGGAAAAGCGCTTTCAGGAGGCCATGGAAATGCT 41


RESULT 9
US-10-027-632-52319
; Sequence 52319, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52319
;   LENGTH: 37
;   TYPE: DNA
```

```
;   ORGANISM: Human
US-10-027-632-52319

  Query Match             1.2%;  Score 20.2;  DB 13;  Length 37;
  Best Local Similarity   75.8%;  Pred. No. 3.8e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        744 ATATCACCAAGGTGATGAGCAAAAGCTCTGCTG 776
              ||||  ||||| ||| |||||||| |||   ||
Db          5 ATATACCCAAGTTGAAGAGCAAAATCTCACGTG 37


RESULT 10
US-10-027-632-52331
; Sequence 52331, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-52331

  Query Match             1.2%;  Score 20.2;  DB 13;  Length 37;
  Best Local Similarity   75.8%;  Pred. No. 3.8e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        744 ATATCACCAAGGTGATGAGCAAAAGCTCTGCTG 776
              ||||  ||||| ||| |||||||| |||   ||
Db          5 ATATACCCAAGTTGAAGAGCAAAATCTCACGTG 37


RESULT 11
US-10-027-632-52319
; Sequence 52319, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52319
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-52319

  Query Match             1.2%;  Score 20.2;  DB 14;  Length 37;
  Best Local Similarity   75.8%;  Pred. No. 3.8e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        744 ATATCACCAAGGTGATGAGCAAAAGCTCTGCTG 776
              ||||  ||||| ||| |||||||| |||   ||
Db          5 ATATACCCAAGTTGAAGAGCAAAATCTCACGTG 37


RESULT 12
US-10-027-632-52331
; Sequence 52331, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-52331

  Query Match             1.2%;  Score 20.2;  DB 14;  Length 37;
  Best Local Similarity   75.8%;  Pred. No. 3.8e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        744 ATATCACCAAGGTGATGAGCAAAAGCTCTGCTG 776
              ||||  ||||| ||| |||||||| |||   ||
Db          5 ATATACCCAAGTTGAAGAGCAAAATCTCACGTG 37


RESULT 13
US-10-006-911-34/c
; Sequence 34, Application US/10006911
; Publication No. US20030125274A1
; GENERAL INFORMATION:
;  APPLICANT: William Gaarde
;  APPLICANT:  Andrew T. Watt
;  TITLE OF INVENTION: ANTISENSE MODULATION OF COLLAPSIN RESPONSE MEDIATOR PROTEIN 2 E
;  FILE REFERENCE: RTS-0200
;  CURRENT APPLICATION NUMBER: US/10/006,911
```

; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 34
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-911-34

Query Match 1.2%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAGGGGAAGAAAAATATTC 20
||||||||||||||||||||
Db 20 TCAGGGGAAGAAAAATATTC 1

RESULT 14
US-10-006-911-35/c
; Sequence 35, Application US/10006911
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: William Gaarde
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF COLLAPSIN RESPONSE MEDIATOR PROTEIN 2 EXP
; FILE REFERENCE: RTS-0200
; CURRENT APPLICATION NUMBER: US/10/006,911
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 35
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-911-35

Query Match 1.2%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AGAAAAATATTCCACGCATC 28
||||||||||||||||||||
Db 20 AGAAAAATATTCCACGCATC 1

RESULT 15
US-10-006-911-36/c
; Sequence 36, Application US/10006911
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: William Gaarde
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF COLLAPSIN RESPONSE MEDIATOR PROTEIN 2 EXP
; FILE REFERENCE: RTS-0200
; CURRENT APPLICATION NUMBER: US/10/006,911
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 36
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-911-36

Query Match 1.2%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 ATATTCCACGCATCACGAGC 34
||||||||||||||||||||
Db 20 ATATTCCACGCATCACGAGC 1

Search completed: December 11, 2003, 00:27:12
Job time : 550 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 10, 2003, 19:07:32 ; Search time 113 Seconds
(without alignments)
6374.666 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976
Perfect score: 1632
Sequence: 1 tcaggggaagaaaaatattc.....ttgtctggtgctcagattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 744296

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	23.2	1.4	39	2	US-08-455-968E-44	Sequence 44, Appl
	2	21.8	1.3	47	4	US-09-641-638-1191	Sequence 1191, Ap
	3	21	1.3	45	1	US-08-440-391-25	Sequence 25, Appl
	4	21	1.3	45	2	US-08-908-597A-25	Sequence 25, Appl
	5	21	1.3	45	3	US-09-236-385A-25	Sequence 25, Appl
	6	21	1.3	45	5	PCT-US96-06122-25	Sequence 25, Appl
c	7	20.6	1.3	41	1	US-08-478-039-40	Sequence 40, Appl
c	8	20.6	1.3	41	1	US-08-476-349A-40	Sequence 40, Appl
c	9	20.6	1.3	41	3	US-08-523-894-33	Sequence 33, Appl
c	10	20.6	1.3	47	4	US-09-422-978-1050	Sequence 1050, Ap
c	11	19.8	1.2	33	3	US-09-054-492B-3	Sequence 3, Appli
c	12	19.8	1.2	41	1	US-08-478-039-39	Sequence 39, Appl
c	13	19.8	1.2	41	1	US-08-476-349A-39	Sequence 39, Appl
c	14	19.8	1.2	41	3	US-08-803-085-5	Sequence 5, Appli
c	15	19.8	1.2	41	3	US-08-523-894-32	Sequence 32, Appl
c	16	19.8	1.2	47	4	US-09-422-978-2072	Sequence 2072, Ap
	17	19.6	1.2	40	4	US-09-472-146A-3	Sequence 3, Appli
c	18	19.6	1.2	47	4	US-09-422-978-3343	Sequence 3343, Ap
	19	19.6	1.2	48	1	US-08-719-331-1	Sequence 1, Appli
	20	19.6	1.2	48	2	US-08-994-719C-1	Sequence 1, Appli
	21	19.6	1.2	50	1	US-08-672-571A-13	Sequence 13, Appl
c	22	19.4	1.2	43	4	US-09-784-130-4	Sequence 4, Appli
c	23	19.2	1.2	45	4	US-09-518-914-11	Sequence 11, Appl
c	24	19.2	1.2	47	4	US-09-641-638-1284	Sequence 1284, Ap
c	25	19.2	1.2	49	3	US-09-275-850-197	Sequence 197, App
c	26	19	1.2	29	4	US-07-672-530C-16	Sequence 16, Appl
c	27	19	1.2	38	4	US-09-438-954-30	Sequence 30, Appl
c	28	19	1.2	47	4	US-09-671-317-950	Sequence 950, App
	29	19	1.2	48	4	US-08-849-567A-98	Sequence 98, Appl
c	30	19	1.2	48	4	US-08-706-945D-70	Sequence 70, Appl
	31	19	1.2	49	1	US-08-015-180-7	Sequence 7, Appli
	32	19	1.2	49	1	US-08-649-196-7	Sequence 7, Appli
	33	19	1.2	49	4	US-09-538-709-1062	Sequence 1062, Ap
c	34	19	1.2	50	3	US-09-315-886C-8	Sequence 8, Appli
c	35	18.8	1.2	31	3	US-09-230-199-19	Sequence 19, Appl
	36	18.8	1.2	34	2	US-08-211-718-14	Sequence 14, Appl
	37	18.8	1.2	45	4	US-09-868-758-16	Sequence 16, Appl
c	38	18.8	1.2	50	3	US-08-957-001B-8	Sequence 8, Appli
c	39	18.8	1.2	50	3	US-09-496-301-8	Sequence 8, Appli
	40	18.6	1.1	33	4	US-09-427-834A-2	Sequence 2, Appli
	41	18.6	1.1	36	2	US-08-418-085-34	Sequence 34, Appl
	42	18.6	1.1	36	3	US-09-099-011A-34	Sequence 34, Appl
c	43	18.6	1.1	47	4	US-09-422-978-2646	Sequence 2646, Ap
	44	18.6	1.1	47	4	US-09-422-978-3871	Sequence 3871, Ap
c	45	18.4	1.1	45	2	US-08-484-993B-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-08-455-968E-44
; Sequence 44, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
US-08-455-968E-44

Query Match 1.4%; Score 23.2; DB 2; Length 39;
Best Local Similarity 77.8%; Pred. No. 2e+03;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy        406 TGTGGACATCAGCGAGTGGCATAAGGGCATCCAGGA 441
              ||||||| ||| | ||  ||| ||| ||||| ||||
Db          4 TGTGGACCTCATCCAGAAGCACAAGAGCATCGAGGA 39
```

```
RESULT 2
US-09-641-638-1191
; Sequence 1191, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 1191
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 10-24-234 : polymorphic base A or G
US-09-641-638-1191

  Query Match             1.3%;  Score 21.8;  DB 4;  Length 47;
  Best Local Similarity   67.4%;  Pred. No. 5.4e+03;
  Matches   29;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        425 CATAAGGGCATCCAGGAGGAGATGGAAGCGCTTGTGAAGGATC 467
              ||  || ||   | | ||||| |: ||| |||| ||||||| |
Db          1 CAGGAGCGCTGTCGGCAGGAGGTRCAAGAGCTTCTGAAGGACC 43


RESULT 3
US-08-440-391-25
; Sequence 25, Application US/08440391
; Patent No. 5656725
;  GENERAL INFORMATION:
;    APPLICANT:  CHITTENDEN, Thomas D.; and
;    APPLICANT:  LUTZ, Robert J.
;    TITLE OF INVENTION:  NOVEL PEPTIDES AND COMPOSITIONS WHICH
;    TITLE OF INVENTION:  MODULATE APOPTOSIS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,391
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, HENRY N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.147
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-942-8400
;      TELEFAX:  202-942-8484
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-440-391-25

  Query Match             1.3%;  Score 21;  DB 1;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 8.9e+03;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | |||||||| |  ||| ||||| ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38


RESULT 4
US-08-908-597A-25
; Sequence 25, Application US/08908597A
; Patent No. 5863795
;  GENERAL INFORMATION:
;    APPLICANT:  CHITTENDEN, Thomas D.; and
;    APPLICANT:  LUTZ, Robert J.
;    TITLE OF INVENTION:  NOVEL PEPTIDES AND COMPOSITIONS WHICH
;    TITLE OF INVENTION:  MODULATE APOPTOSIS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/908,597A
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,391
;      FILING DATE:  12-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, HENRY N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.147
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-942-8400
;      TELEFAX:  202-942-8484
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-908-597A-25

  Query Match             1.3%;  Score 21;  DB 2;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 8.9e+03;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | |||||||| |  ||| ||||| ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38
```

```
RESULT 5
US-09-236-385A-25
; Sequence 25, Application US/09236385A
; Patent No. 6221615
;  GENERAL INFORMATION:
;       APPLICANT: CHITTENDEN, Thomas D.; and
;                  LUTZ, Robert J.
;       TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
;                           MODULATE APOPTOSIS
;       NUMBER OF SEQUENCES: 41
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Hale and Dorr
;            STREET: 1455 Pennsylvania Avenue, N.W.
;            CITY: Washington
;            STATE: D.C.
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/236,385A
;            FILING DATE: 25-Jan-1999
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: WIXON, HENRY N.
;            REGISTRATION NUMBER: 32,073
;                                 (C) ATTORNEY DOCKET NO. 104322.147CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 202-942-8400
;            TELEFAX: 202-942-8484
;   INFORMATION FOR SEQ ID NO: 25:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 45 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-236-385A-25

  Query Match             1.3%;  Score 21;  DB 3;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 8.9e+03;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | |||||||| |  ||| |||||  ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38


RESULT 6
PCT-US96-06122-25
; Sequence 25, Application PC/TUS9606122
;  GENERAL INFORMATION:
;    APPLICANT:  IMMUNOGEN, INC.
;    TITLE OF INVENTION:  NOVEL PEPTIDES AND COMPOSITIONS
;    TITLE OF INVENTION:  WHICH MODULATE APOPTOSIS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hale and Dorr
;      STREET:  1455 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/06122
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/440,391
;      FILING DATE:  12-MAY-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WIXON, HENRY N.
;      REGISTRATION NUMBER:  32,073
;      REFERENCE/DOCKET NUMBER:  104322.147PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-942-8400
;      TELEFAX:  202-942-8484
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US96-06122-25

  Query Match             1.3%;  Score 21;  DB 5;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 8.9e+03;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | |||||||| |  ||| |||||  ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38


RESULT 7
US-08-478-039-40/c
; Sequence 40, Application US/08478039
; Patent No. 5681722
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,039
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-160
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens or Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  kappa light chain primer with BglII site
US-08-478-039-40

  Query Match             1.3%;  Score 20.6;  DB 1;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.1e+04;
  Matches   27;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
             ||||| ||||  |||:||| ||    |||  || |||||
Db         41 CTGAGTGGGGGCCCTBGTGTCCATGGTGAGAGATCTGTG 3


RESULT 8
US-08-476-349A-40/c
; Sequence 40, Application US/08476349A
; Patent No. 5750105
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,349A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-161
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens or Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  kappa light chain primer with BglII site
US-08-476-349A-40

  Query Match             1.3%;  Score 20.6;  DB 1;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.1e+04;
  Matches   27;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
             ||||| ||||  |||:||| ||    |||  || |||||
Db         41 CTGAGTGGGGGCCCTBGTGTCCATGGTGAGAGATCTGTG 3


RESULT 9
US-08-523-894-33/c
; Sequence 33, Application US/08523894
; Patent No. 6136310
;  GENERAL INFORMATION:
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Reff, Mitchell E.
;    TITLE OF INVENTION:  Recombinant Anti-CD4 Antibodies for Human
;    TITLE OF INVENTION:  Therapy
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince Street
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22314-3187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/523,894
;      FILING DATE:  06-SEP-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin, Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-165
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Human or Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  kappa light chain primer with Bgl II site
US-08-523-894-33

  Query Match             1.3%;  Score 20.6;  DB 3;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.1e+04;
  Matches   27;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;
```

```
Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
              ||||| ||||  |||:||| ||    |||  || |||||
Db         41 CTGAGTGGGGGCCCTBGTGTCCATGGTGAGAGATCTGTG 3


RESULT 10
US-09-422-978-1050/c
; Sequence 1050, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1050
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-20023-386 : polymorphic base A or T
US-09-422-978-1050

  Query Match             1.3%;  Score 20.6;  DB 4;  Length 47;
  Best Local Similarity   64.4%;  Pred. No. 1.2e+04;
  Matches   29;  Conservative    1;  Mismatches   15;  Indels    0;  Gaps    0;

Qy        477 ATTCCTTCCTCGTGTACATGGCTTTCAAAGATCGCTTCCAGCTAA 521
              | | ||||||| ||||||||||:||    ||   |   ||| | |
Db      '  46 AATGCTTCCTCTTGTACATGGCWTTTCTGGACTACAAGCAGGTCA 2


RESULT 11
US-09-054-492B-3/c
; Sequence 3, Application US/09054492B
; Patent No. 6218115
;  GENERAL INFORMATION:
;    APPLICANT:  TAKESHI NAKAMURA
;    TITLE OF INVENTION:  HUMAN CYCLIN I AND GENES ENCODING SAME
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PILLSBURY, MADISON & SUTRO
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/054,492B
;      FILING DATE:  APRIL 3, 1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PAUL E. WHITE, JR.
;      REGISTRATION NUMBER:  32,011
;      REFERENCE/DOCKET NUMBER:  7898/252159
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627CUSH
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-054-492B-3

  Query Match             1.2%;  Score 19.8;  DB 3;  Length 33;
  Best Local Similarity   77.4%;  Pred. No. 1.7e+04;
  Matches   24;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1377 CCGAAGGCTCTGGACGCTACATTCCCCGGAA 1407
              || ||||| ||||| || ||| ||| ||||
Db         33 CCAAAGGCCCTGGAAACTTCATACCCGGGAA 3


RESULT 12
US-08-478-039-39/c
; Sequence 39, Application US/08478039
; Patent No. 5681722
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,039
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-160
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ANTI-SENSE:  NO
```

```
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens and Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  kappa light chain primer with BglII site
US-08-478-039-39

  Query Match             1.2%;  Score 19.8;  DB 1;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.9e+04;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
                ||||| ||||  |||| || ||    |||  || |||||
Db         41 CTGAGCGGGGACCCTCATGTCCATGGTGAGAGATCTGTG 3


RESULT 13
US-08-476-349A-39/c
; Sequence 39, Application US/08476349A
; Patent No. 5750105
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,349A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-161
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens and Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  kappa light chain primer with BglII site
US-08-476-349A-39

  Query Match             1.2%;  Score 19.8;  DB 1;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.9e+04;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
                ||||| ||||  |||| || ||    |||  || |||||
Db         41 CTGAGCGGGGACCCTCATGTCCATGGTGAGAGATCTGTG 3


RESULT 14
US-08-803-085-5/c
; Sequence 5, Application US/08803085
; Patent No. 6011138
;  GENERAL INFORMATION:
;    APPLICANT:  REFF, Mitchell E.
;    APPLICANT:  KLOETZER, William S.
;    APPLICANT:  NAKAMURA, Takehiko
;    TITLE OF INVENTION:  GAMMA-1 ANTI-HUMAN CD23 MONOCLONAL
;    TITLE OF INVENTION:  ANTIBODIES AND USE THEREOF AS THERAPEUTICS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/803,085
;      FILING DATE:  20-FEB-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin, Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-353
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-803-085-5

  Query Match             1.2%;  Score 19.8;  DB 3;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.9e+04;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
                ||||| ||||  |||| || ||    |||  || |||||
Db         41 CTGAGCGGGGACCCTCATGTCCATGGTGAGAGATCTGTG 3


RESULT 15
US-08-523-894-32/c
; Sequence 32, Application US/08523894
; Patent No. 6136310
;  GENERAL INFORMATION:
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Reff, Mitchell E.
;    TITLE OF INVENTION:  Recombinant Anti-CD4 Antibodies for Human
;    TITLE OF INVENTION:  Therapy
;    NUMBER OF SEQUENCES:  59
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince Street
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22314-3187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/523,894
;      FILING DATE:  06-SEP-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin, Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-165
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Human or Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  kappa light chain primer with Bgl II site
US-08-523-894-32

  Query Match             1.2%;  Score 19.8;  DB 3;  Length 41;
  Best Local Similarity   69.2%;  Pred. No. 1.9e+04;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1463 CTGAGAGGGGTTCCTCGTGGCCTGTATGACGGACCTGTG 1501
              ||||| ||||  |||| || ||    |||  || |||||
Db         41 CTGAGCGGGGACCCTCATGTCCATGGTGAGAGATCTGTG 3


Search completed: December 10, 2003, 22:41:01
Job time : 116 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 10, 2003, 15:49:05 ; Search time 481 Seconds
(without alignments)
9159.008 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976
Perfect score: 1632
Sequence: 1 tcaggggaagaaaaatattc.....ttgtctggtgctcagattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 2467832

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	31	1.9	31	22	AAI29604	Human single nucle
	2	31	1.9	31	22	AAI29605	Human single nucle
	3	23.2	1.4	39	20	AAX02133	Human FEN-1 DNA fr
	4	23	1.4	41	24	ABK14176	Insulin like growt
c	5	22	1.3	22	19	AAV60829	Primer 509 for ULI
c	6	22	1.3	22	24	AAD29279	Human CRMP-2 ampli
c	7	21.8	1.3	50	22	AAL32108	Human SNP oligonuc
	8	21.4	1.3	41	24	ABK14175	Insulin like growt
	9	21	1.3	21	22	AAF95365	Human gene single
	10	21	1.3	21	22	AAF95366	Human gene single
	11	21	1.3	21	22	AAF95367	Human gene single
	12	21	1.3	21	24	AAD29278	Human CRMP-2 ampli
	13	21	1.3	45	17	AAT42432	GD domain region f
	14	21	1.3	45	24	ABK11181	DNA encoding Bax G
c	15	21	1.3	47	21	AAZ66703	Human map-related
	16	20.8	1.3	50	22	AAL28266	Human SNP oligonuc
c	17	20.6	1.3	41	14	AAQ35919	Human/monkey kappa
c	18	20.6	1.3	41	18	AAT95144	Human or monkey Ig
c	19	20.6	1.3	41	18	AAT92220	Monkey/human kappa
c	20	20.6	1.3	41	18	AAT62891	Human or monkey ka
c	21	20.6	1.3	41	19	AAV31404	Human or monkey ka
c	22	20.6	1.3	41	19	AAV23782	Primer for Anti-CD
c	23	20.6	1.3	41	19	AAV05666	Human/monkey kappa
c	24	20.6	1.3	41	25	ABX76660	Immunoglobulin gen
	25	20.6	1.3	47	22	AAF89202	Sample member clus
	26	20.6	1.3	50	22	AAL30277	Human SNP oligonuc
	27	20.6	1.3	50	24	ABK91333	Human JFY1, p53-bi
c	28	20.4	1.2	33	24	ABZ20637	Human ribonucleoti
c	29	20.4	1.2	40	25	ABZ24571	Stat1 SH2 PCR prim
	30	20.4	1.2	41	25	ACC42038	Human SCN5A gene r
c	31	20.2	1.2	50	24	ABK91334	Human JFY1, p53-bi
	32	20	1.2	20	19	AAV60828	Primer 5227 for UL
c	33	20	1.2	49	21	AAA75784	PCR primer for a h
c	34	20	1.2	50	25	ABZ68287	Nucleotide sequenc
c	35	19.8	1.2	33	18	AAT73938	Primer for human c
c	36	19.8	1.2	40	25	ACA55269	Rat IgE CH3 domain
c	37	19.8	1.2	41	14	AAQ35918	Human/monkey kappa
c	38	19.8	1.2	41	18	AAT95143	Human or monkey Ig
c	39	19.8	1.2	41	18	AAT92219	Monkey/human kappa
c	40	19.8	1.2	41	18	AAT62890	Human or monkey ka
c	41	19.8	1.2	41	19	AAV33311	Anti-CD23 5E8 anti
c	42	19.8	1.2	41	19	AAV31403	Human or monkey ka
c	43	19.8	1.2	41	19	AAV23781	Primer for Anti-CD
c	44	19.8	1.2	41	19	AAV05665	Human/monkey kappa
c	45	19.8	1.2	41	25	ABX76659	Immunoglobulin gen

ALIGNMENTS

RESULT 1
AAI29604
ID AAI29604 standard; DNA; 31 BP.
XX
AC AAI29604;
XX
DT 18-OCT-2001 (first entry)
XX
DE Human single nucleotide polymorphism (SNP) DPYSL2 1.
XX
KW Human; resequence; genotype; disease; forensic; paternity testing;
KW single nucleotide polymorphism; SNP; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Variation replace(16,A)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200166800-A2.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001; 2001WO-US07268.
XX
PR 07-MAR-2000; 2000US-0187510.
PR 22-MAY-2000; 2000US-0206129.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX

PI Cargill M, Ireland JS, Lander ES;
XX
DR WPI; 2001-522952/57.
XX
PT Nucleic acid molecules from the human genome which include polymorphic
PT sites, useful in methods for predicting the presence, absence or
PT severity of a particular phenotype or disorder (e.g. diabetes)
PT associated with a particular genotype -
XX
PS Claim 1; Page 34; 145pp; English.
XX
CC The invention relates to the identification of nucleic acid molecules
CC (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC which can predispose individuals to disease. Various genes from a number
CC of individuals were resequenced and single nucleotide polymorphisms
CC (SNPs) in these genes discovered. The method is useful for predicting the
CC presence, absence or severity of a particular phenotype or disorder (e.g.
CC diabetes) associated with a particular genotype. The nucleic acids
CC containing the polymorphic sites may be useful in forensics and paternity
CC testing.
XX
SQ Sequence 31 BP; 5 A; 10 C; 9 G; 7 T; 0 other;

Query Match 1.9%; Score 31; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 GGGACAAGCCTGCTCGCTGCCTTTGACCAGT 359
||||||||||||||||||||||||||||||
Db 1 GGGACAAGCCTGCTCGCTGCCTTTGACCAGT 31

RESULT 2
AAI29605
ID AAI29605 standard; DNA; 31 BP.
XX
AC AAI29605;
XX
DT 18-OCT-2001 (first entry)
XX
DE Human single nucleotide polymorphism (SNP) DPYSL2 2.
XX
KW Human; resequence; genotype; disease; forensic; paternity testing;
KW single nucleotide polymorphism; SNP; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Variation replace(16,G)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200166800-A2.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001; 2001WO-US07268.
XX
PR 07-MAR-2000; 2000US-0187510.
PR 22-MAY-2000; 2000US-0206129.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Cargill M, Ireland JS, Lander ES;
XX
DR WPI; 2001-522952/57.
XX
PT Nucleic acid molecules from the human genome which include polymorphic
PT sites, useful in methods for predicting the presence, absence or
PT severity of a particular phenotype or disorder (e.g. diabetes)
PT associated with a particular genotype -
XX
PS Claim 1; Page 34; 145pp; English.
XX
CC The invention relates to the identification of nucleic acid molecules
CC (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC which can predispose individuals to disease. Various genes from a number
CC of individuals were resequenced and single nucleotide polymorphisms
CC (SNPs) in these genes discovered. The method is useful for predicting the
CC presence, absence or severity of a particular phenotype or disorder (e.g.
CC diabetes) associated with a particular genotype. The nucleic acids
CC containing the polymorphic sites may be useful in forensics and paternity
CC testing.
XX
SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 other;

Query Match 1.9%; Score 31; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 973 GGTCACGGGCAGTGCCCATTGCACGTTTAAC 1003
||||||||||||||||||||||||||||||
Db 1 GGTCACGGGCAGTGCCCATTGCACGTTTAAC 31

RESULT 3
AAX02133
ID AAX02133 standard; DNA; 39 BP.
XX
AC AAX02133;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 DNA fragment #16.
XX
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplasmic agent; cleavage; ss.
XX
OS Homo sapiens.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI; 1999-179985/15.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
XX
PS Claim 2; Column 25-26; 58pp; English.
XX
CC AAX02118-X02140 represent fragments of a human FEN-1 (flap endonuclease)
CC protein. This protein can be used in methods for detecting a
CC pathological condition in a patient, for diagnostic purposes, for
CC screening for antineoplasmic agents and carcinogens, for diagnostic
CC staging of neoplasma, for producing recombinant flap endonuclease for use
CC as research or diagnostic reagents, for producing antibodies reactive
CC with the novel polypeptides, for producing transgenic nonhuman animals
CC expressing the novel polypeptides encoded by a transgene. The invention
CC also provides novel molecular cloning techniques and reagents involving
CC cleavage of a flap or nick with a flap endonuclease.
XX
SQ Sequence 39 BP; 12 A; 10 C; 12 G; 5 T; 0 other;

Query Match 1.4%; Score 23.2; DB 20; Length 39;
Best Local Similarity 77.8%; Pred. No. 1.2e+04;

```
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        406 TGTGGACATCAGCGAGTGGCATAAGGGCATCCAGGA 441
              ||||||| ||| | ||  ||| ||| ||||| ||||
Db          4 TGTGGACCTCATCCAGAAGCACAAGAGCATCGAGGA 39


RESULT 4
ABK14176
ID   ABK14176 standard; DNA; 41 BP.
XX
AC   ABK14176;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Insulin like growth factor binding protein 11.88 cDNA probe #2.
XX
KW   Insulin like growth factor binding protein 11.88; probe; ss; cytostatic;
KW   embryo development dysmorphia; malignant tumour; gene therapy; cancer.
XX
OS   Unidentified.
XX
PN   WO200212493-A1.
XX
PD   14-FEB-2002.
XX
PF   11-JUN-2001; 2001WO-CN00951.
XX
PR   14-JUN-2000; 2000CN-0116491.
XX
PA   (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-172159/22.
XX
PT   Insulin like growth factor binding protein 11.88 and encoding
PT   polynucleotide, used in diagnosis and treatment of malignant tumours -
XX
PS   Example 6; Page 16; 38pp; Chinese.
XX
CC   The invention relates to an insulin like growth factor binding protein
CC   11.88 and the DNA sequence encoding the polypeptide. The DNA and protein
CC   are used in diagnosis and treatment of malignant tumour and dysmorphia of
CC   development of an embryo. This sequence represents a probe which
CC   hybridises to cDNA which encodes the insulin like growth factor binding
CC   protein 11.88 of the invention.
XX
SQ   Sequence 41 BP; 4 A; 15 C; 11 G; 11 T; 0 other;

  Query Match             1.4%;  Score 23;  DB 24;  Length 41;
  Best Local Similarity   74.4%;  Pred. No. 1.4e+04;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1195 TGCTGTGGGATCCGATGCCGACCTGGTCATCTGGGACCC 1233
              ||||| |||||||  | ||  ||  ||||| |||||| |
Db          3 TGCTGGGGGATCCCTTCCCCCCCCAGTCATTTGGGACTC 41


RESULT 5
AAV60829/c
ID   AAV60829 standard; DNA; 22 BP.
XX
AC   AAV60829;
XX
DT   08-DEC-1998  (first entry)
XX
DE   Primer 509 for ULIP-2 gene.
XX
KW   Mouse; Unc-33-like phospho-protein; ULIP; rat; neoplasm; tumourigenesis;
KW   neurodegenerative disorder; diagnosis; primer; PCR; amplification; ss.
XX
OS   Synthetic.
OS   Mus musculus.
XX
PN   FR2759701-A1.
XX
PD   21-AUG-1998.
XX
PF   19-FEB-1997;   97FR-0001961.
XX
PR   19-FEB-1997;   97FR-0001961.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aguera M,  Belin MF,  Byk T,  Honnorat J,  Kolattukudy P;
PI   Quach TT,  Sobel A;
XX
DR   WPI; 1998-449610/39.
XX
PT   Mouse and human ULIP poly:peptide(s) - useful in detection of
PT   para-neoplastic neurological syndromes
XX
PS   Example 10; Page 43; 90pp; French.
XX
CC   Primers AAV60828-V60829 represent PCR primers used to amplify the gene
CC   sequence of the Unc-33-like phosphoprotein 2 (ULIP-2) gene from lung
CC   cancer cells, especially from small cell lung cancers.  The
CC   amplification is used to detect the ULIP-3 gene is these cancers.
CC   Proteins of the ULIP family or their corresponding nucleic acids can be
CC   used in compositions for treating neurodegenerative disorders and
CC   neoplasms, especially for para-neoplastic neurological syndromes and/or
CC   for the early diagnosis of tumourigenesis.
XX
SQ   Sequence 22 BP; 3 A; 8 C; 4 G; 7 T; 0 other;

  Query Match             1.3%;  Score 22;  DB 19;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        409 GGACATCAGCGAGTGGCATAAG 430
              ||||||||||||||||||||||
Db         22 GGACATCAGCGAGTGGCATAAG 1


RESULT 6
AAD29279/c
ID   AAD29279 standard; DNA; 22 BP.
XX
AC   AAD29279;
XX
DT   07-MAY-2002  (first entry)
XX
DE   Human CRMP-2 amplifying PCR primer, CRMP2-4.
XX
KW   Human; collapsin response-mediator protein-5; CRMP-5; neoplasm;
KW   paraneoplastic autoimmunity; small-cell lung carcinoma; thymoma;
KW   neuroblastoma; CRMP-2; dihydropyrimidinase related protein-2;
KW   DRP-2; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200202620-A2.
XX
PD   10-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US20507.
XX
PR   29-JUN-2000; 2000US-0606924.
XX
PA   (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
XX
PI   Lennon VA,  Yu Z,  Kryzer TJ,  Griesmann GE;
XX
DR   WPI; 2002-171637/22.
XX
```

```
PT   Isolated nucleic acids (N) which encode CRMP-5 polypeptides are useful
PT   in detecting anti-CRMP-5 autoantibodies in a patient with
PT   paraneoplastic neurological manifestations and neoplasm -
XX
PS   Example 2; Page 9; 40pp; English.
XX
CC   The invention relates to human collapsin response-mediator protein-5
CC   (CRMP-5) and nucleic acid molecules encoding such proteins. The
CC   presence of anti-CRMP-5 autoantibody in a biological sample is
CC   associated with paraneoplastic autoimmunity in the individual and
CC   neoplasm such as small-cell lung carcinoma, neuroblastoma and thymoma.
CC   The present sequence is human CRMP-2 (dihydropyrimidinase related
CC   protein-2, DRP-2) amplifying PCR primer used in the exemplification
CC   of the invention.
XX
SQ   Sequence 22 BP; 2 A; 4 C; 8 G; 8 T; 0 other;

  Query Match             1.3%;  Score 22;  DB 24;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1229 GACCCCGACAGCGTTAAAACCA 1250
              ||||||||||||||||||||||
Db         22 GACCCCGACAGCGTTAAAACCA 1

RESULT 7
AAL32108/c
ID   AAL32108 standard; DNA; 50 BP.
XX
AC   AAL32108;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #5316.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2917; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 12 A; 22 C; 8 G; 8 T; 0 other;

  Query Match             1.3%;  Score 21.8;  DB 22;  Length 50;
  Best Local Similarity   70.7%;  Pred. No. 3.3e+04;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1454 CTGGCTGAGCTGAGAGGGGTTCCTCGTGGCCTGTATGACGG 1494
              |||| || |   ||| ||| |||| |||||| || || ||
Db         50 CTGGGTGGGTATAGAAGGGCTCCTGCTGGCCTCTAGGATGG 10

RESULT 8
ABK14175
ID   ABK14175 standard; DNA; 41 BP.
XX
AC   ABK14175;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Insulin like growth factor binding protein 11.88 cDNA probe #1.
XX
KW   Insulin like growth factor binding protein 11.88; probe; ss; cytostatic;
KW   embryo development dysmorphia; malignant tumour; gene therapy; cancer.
XX
OS   Unidentified.
XX
PN   WO200212493-A1.
XX
PD   14-FEB-2002.
XX
PF   11-JUN-2001; 2001WO-CN00951.
XX
PR   14-JUN-2000; 2000CN-0116491.
XX
PA   (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-172159/22.
XX
PT   Insulin like growth factor binding protein 11.88 and encoding
PT   polynucleotide, used in diagnosis and treatment of malignant tumours -
XX
PS   Example 6; Page 15; 38pp; Chinese.
XX
CC   The invention relates to an insulin like growth factor binding protein
CC   11.88 and the DNA sequence encoding the polypeptide. The DNA and protein
CC   are used in diagnosis and treatment of malignant tumour and dysmorphia of
CC   development of an embryo. This sequence represents a probe which
CC   hybridises to cDNA which encodes the insulin like growth factor binding
CC   protein 11.88 of the invention.
XX
SQ   Sequence 41 BP; 5 A; 14 C; 11 G; 11 T; 0 other;

  Query Match             1.3%;  Score 21.4;  DB 24;  Length 41;
  Best Local Similarity   71.8%;  Pred. No. 3.9e+04;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1195 TGCTGTGGGATCCGATGCCGACCTGGTCATCTGGGACCC 1233
              ||||| |||||||  | |   ||  ||||| ||||||
Db          3 TGCTGGGGGATCCCTTCCACCCCCAGTCATTTGGGACTC 41
```

```
RESULT 9
AAF95365
ID   AAF95365 standard; DNA; 21 BP.
XX
AC   AAF95365;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #126.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,G)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 57; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 7 A; 4 C; 9 G; 1 T; 0 other;

  Query Match             1.3%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        608 GCAGAGGAGCAGCAGAGGATC 628
              |||||||||||||||||||||
Db          1 GCAGAGGAGCAGCAGAGGATC 21


RESULT 10
AAF95366
ID   AAF95366 standard; DNA; 21 BP.
XX
AC   AAF95366;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #127.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 57; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 5 A; 3 C; 8 G; 5 T; 0 other;

  Query Match             1.3%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1488 ATGACGGACCTGTGTGTGAAG 1508
              |||||||||||||||||||||
Db          1 ATGACGGACCTGTGTGTGAAG 21


RESULT 11
AAF95367
ID   AAF95367 standard; DNA; 21 BP.
XX
AC   AAF95367;
XX
```

```
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #128.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 57; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 7 A; 10 C; 2 G; 2 T; 0 other;

  Query Match             1.3%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        708 CCATCACCATCGCCAACCAGA 728
              |||||||||||||||||||||
Db          1 CCATCACCATCGCCAACCAGA 21


RESULT 12
AAD29278
ID   AAD29278 standard; DNA; 21 BP.
XX
AC   AAD29278;
XX
DT   07-MAY-2002  (first entry)
XX
DE   Human CRMP-2 amplifying PCR primer, CRMP2-3.
XX
KW   Human; collapsin response-mediator protein-5; CRMP-5; neoplasm;
KW   paraneoplastic autoimmunity; small-cell lung carcinoma; thymoma;
KW   neuroblastoma; CRMP-2; dihydropyrimidinase related protein-2;
KW   DRP-2; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200202620-A2.
XX
PD   10-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US20507.
XX
PR   29-JUN-2000; 2000US-0606924.
XX
PA   (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
XX
PI   Lennon VA,  Yu Z,  Kryzer TJ,  Griesmann GE;
XX
DR   WPI; 2002-171637/22.
XX
PT   Isolated nucleic acids (N) which encode CRMP-5 polypeptides are useful
PT   in detecting anti-CRMP-5 autoantibodies in a patient with
PT   paraneoplastic neurological manifestations and neoplasm  -
XX
PS   Example 2; Page 9; 40pp; English.
XX
CC   The invention relates to human collapsin response-mediator protein-5
CC   (CRMP-5) and nucleic acid molecules encoding such proteins. The
CC   presence of anti-CRMP-5 autoantibody in a biological sample is
CC   associated with paraneoplastic autoimmunity in the individual and
CC   neoplasm such as small-cell lung carcinoma, neuroblastoma and thymoma.
CC   The present sequence is human CRMP-2 (dihydropyrimidinase related
CC   protein-2, DRP-2) amplifying PCR primer used in the exemplification
CC   of the invention.
XX
SQ   Sequence 21 BP; 6 A; 6 C; 5 G; 4 T; 0 other;

  Query Match             1.3%;  Score 21;  DB 24;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+04;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        505 AGATCGCTTCCAGCTAACGGA 525
              |||||||||||||||||||||
Db          1 AGATCGCTTCCAGCTAACGGA 21


RESULT 13
AAT42432
ID   AAT42432 standard; DNA; 45 BP.
XX
AC   AAT42432;
XX
DT   29-JUL-1997  (first entry)
XX
DE   GD domain region for Bax encoding DNA, amino acid residues 59-73.
XX
KW   Apoptosis; follicular lymphoma; tumour; p53; antibody; ss.
XX
OS   Synthetic.
XX
PN   WO9635951-A1.
XX
PD   14-NOV-1996.
XX
PF   06-MAY-1996;   96WO-US06122.
XX
PR   12-MAY-1995;   95US-0440391.
XX
PA   (IMMU-) IMMUNOGEN INC.
XX
PI   Chittenden TD,  Lutz RJ;
XX
```

```
DR   WPI; 1996-518805/51.
DR   P-PSDB; AAW06296.
XX
PT   Peptide(s) comprising GD domains - have similar activities to wild
PT   type Bak, and cause cellular apoptosis for treatment of viral
PT   infection
XX
PS   Claim 6; Page 47; 69pp; English.
XX
CC   The term GD domain refers to a protein domain first identified in
CC   Bak and shown to be essential for the interaction of Bak with Bcl-x(L)
CC   and for Bak's cell killing function; and to peptides and/or molecules
CC   capable of mimicking its structure and/or function. The present sequence
CC   encodes a GD domain corresponding to amino acid residues 59-73 of Bax.
CC   An antibody raised against a GD domain may be used to screen a cDNA
CC   expression library for clones comprising cDNA inserts encoding
CC   immunocrossreactive proteins. Truncated GD domain peptides have been
CC   shown to maintain the protein binding and cell killing function
CC   exhibited by wild type Bak. These molecules may induce apoptosis in
CC   tumour cell. These peptides act independently of p53 status. Bak or
CC   GD domain mimetics that inhibit Bcl-2 may be selectively toxic to
CC   certain tumours, e.g. follicular lymphoma, which depend on high levels
CC   of Bcl-2 for their continued growth and survival. GD domain mimetics
CC   may also be used for combatting viral infections by causing apoptosis
CC   of infected cells.
XX
SQ   Sequence 45 BP; 12 A; 11 C; 15 G; 7 T; 0 other;

  Query Match             1.3%;  Score 21;  DB 17;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 5.2e+04;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | ||||||||  |  ||| |||||  ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38


RESULT 14
ABK11181
ID   ABK11181 standard; DNA; 45 BP.
XX
AC   ABK11181;
XX
DT   05-JUN-2002  (first entry)
XX
DE   DNA encoding Bax GD domain region #1 for modulating apoptosis.
XX
KW   GD domain; apoptosis; interaction with Bcl-XL; cell killing function;
KW   bak; cell death regulatory molecule; autoimmune disease; cancer;
KW   bax; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..45
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Bax GD domain region #1"
FT                   /note= "This sequence lacks both start and stop codons"
XX
PN   US6221615-B1.
XX
PD   24-APR-2001.
XX
PF   25-JAN-1999;   99US-0236385.
XX
PR   12-MAY-1995;   95US-0440391.
PR   08-AUG-1997;   97US-0908597.
XX
PA   (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI   Chittenden TD,  Lutz RJ;
XX
DR   WPI; 2002-234950/29.
DR   P-PSDB; AAU77879.
XX
PT   Identifying agents (e.g. modulators of apoptosis) capable of modulating
PT   GD domain mediated heterodimerisation or homodimerisation comprises
PT   carrying out a heterodimerisation or homodimerisation assay  -
XX
PS   Disclosure; Fig 8B; 37pp; English.
XX
CC   The present invention relates to novel peptides, designated GD domains,
CC   which are capable of modulating apoptosis. The GD domains are essential
CC   for Bak's interaction with Bcl-XL, and to Bak's cell killing function.
CC   The GD domains mediate key protein/protein interactions with multiple
CC   cell death regulatory molecules. Also described are methods of
CC   identifying agonists or antagonists of GD domains. The methods are
CC   useful for identifying agents capable of modulating GD domain mediated
CC   heterodimerisation or homodimerisation. The methods are particularly
CC   useful in drug screening and design, e.g. for identifying agents for
CC   treating autoimmune disease or cancer, or for identifying modulators
CC   of apoptosis. The present DNA sequence encodes a GD domain region.
XX
SQ   Sequence 45 BP; 12 A; 11 C; 15 G; 7 T; 0 other;

  Query Match             1.3%;  Score 21;  DB 24;  Length 45;
  Best Local Similarity   73.0%;  Pred. No. 5.2e+04;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        414 TCAGCGAGTGGCATAAGGGCATCCAGGAGGAGATGGA 450
              | ||||||||  |  ||| |||||  ||| ||  ||||
Db          2 TGAGCGAGTGTCTCAAGCGCATCGGGGACGAACTGGA 38


RESULT 15
AAZ66703/c
ID   AAZ66703 standard; DNA; 47 BP.
XX
AC   AAZ66703;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1050.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 1; Page 454; 2745pp; English.
```

XX
CC AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the
CC invention have a variety of uses: they can be used for high density
CC mapping of the human genome, and in complex association studies and
CC haplotyping studies which are useful in determining the genetic basis
CC for disease states. Compositions and methods of the invention can also
CC be useful for the identification of the targets for the development of
CC pharmaceutical agents and diagnostic methods, as well as the
CC characterisation of the differential efficacious responses to and side
CC effects from pharmaceutical agents acting on a disease as well as other
CC treatment.
CC N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC and 3367, are not actually given a sequence in the Sequence Listing
CC from the present invention.
XX
SQ Sequence 47 BP; 15 A; 11 C; 11 G; 10 T; 0 other;

Query Match 1.3%; Score 21; DB 21; Length 47;
Best Local Similarity 66.7%; Pred. No. 5.4e+04;
Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy        477 ATTCCTTCCTCGTGTACATGGCTTTCAAAGATCGCTTCCAGCTAA 521
              | | ||||||| ||||||||||||    ||   |   ||| | |
Db         46 AATGCTTCCTCTTGTACATGGCTTTTCTGGACTACAAGCAGGTCA 2
```

Search completed: December 10, 2003, 19:47:32
Job time : 483 secs